Real-Time Acquisition of High Quality Face Sequences from an Active Pan-Tilt-Zoom Camera

Haque, Mohammad A.; Nasrollahi, Kamal; Moeslund, Thomas B.

Published in:
2013 10th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)

DOI (link to publication from Publisher):
10.1109/AVSS.2013.6636680

Publication date:
2013

Document Version
Accepted author manuscript, peer reviewed version

Real-Time Acquisition of High Quality Face Sequences from an Active Pan-Tilt-Zoom Camera

Mohammad A. Haque, Kamal Nasrollahi, Thomas B. Moeslund
Visual Analysis of People Laboratory, Aalborg University (AAU), Aalborg, Denmark
{mah, kn, tbm}@create.aau.dk

Abstract

Traditional still camera-based facial image acquisition systems in surveillance applications produce low quality face images. This is mainly due to the distance between the camera and subjects of interest. Furthermore, people in such videos usually move around, change their head poses, and facial expressions. Moreover, the imaging conditions like illumination, occlusion, and noise may change. These all aggregate the quality of most of the detected face images in terms of measures like resolution, pose, brightness, and sharpness. To deal with these problems this paper presents an active camera-based real-time high-quality face image acquisition system, which utilizes pan-tilt-zoom parameters of a camera to focus on a human face in a scene and employs a face quality assessment method to log the best quality faces from the captured frames. The system consists of four modules: face detection, camera control, face tracking, and face quality assessment before logging. Experimental results show that the proposed system can effectively log the high quality faces from the active camera in real-time (an average of 61.74ms was spent per frame) with an accuracy of 85.27% compared to human annotated data.

1. Introduction

Computer-based automatic analysis of facial image has important applications in many different areas, such as, surveillance, medical diagnosis, biometrics, expression recognition and social cue analysis [1]. However, acquisition of qualified and applicable images from a camera setup or a video clip is essentially the first step of facial image analysis. The application performance greatly depends upon the quality of the face in the image. A number of parameters such as face resolution, pose, brightness, and sharpness determine the quality of a face image. When a video based practical image acquisition system produces too many facial images to be processed in surveillance applications, most of these images are subjected to the aforementioned problems [2]. For example, a human face at 5 meters distance from the camera subtends only about 4x6 pixels on a 640x480 sensor with 130 degrees field of view, which is mostly insufficient resolution for further processing [3]. In order to get rid of the computational cost and inaccuracy problem of low quality facial image processing, a Face Quality Assessment (FQA) technique can be employed to select the qualified faces from the image frames captured by a camera [4]. This reduces significant amount of disqualified faces from the captured imagery and keeps a small number of best face images which are named as face sequence or so called Face Log. However, FQA merely checks whether a face is qualified or not, and cannot increase the quality of captured face. Besides, processing time to determine the face quality is also an obstacle to achieve real-time performance while extracting faces from video sequences. A possible solution to low resolution faces from a video sequence is to employ techniques like Super-Resolution (SR) image reconstruction algorithms, to generate high resolution faces [5]. However, it is still an active research area and subjected to problems like huge computational costs, uncertainty in reconstructed high frequency components, and small doable magnification factor [6]. On the other hand, due to improvements in digital Pan-Tilt-Zoom (PTZ) camera technology and availability of PTZ cameras in low cost, acquisition of high resolution facial images in surveillance, forensic and medical applications by directly using camera instead of SR techniques is an emerging field of study [7].

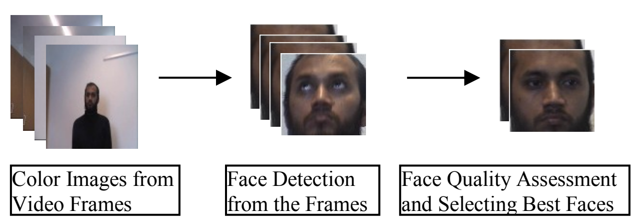

Figure 1. Block diagram of a typical facial image acquisition system with face quality measure

A typical facial image acquisition system including FQA is shown in Figure 1, which was proposed in [4]. Given a video sequence in the face detection step the regions including a face are detected and are considered as Region of Interest (ROI) in the frames. Then, a FQA algorithm assesses the ROIs in terms of face quality measures and stores the best quality faces in face log.

A vast body of literature addressed different issues of facial image acquisition. An appearance-based approach to real-time face detection from video frames was proposed in [8] which merely work in real-time in VGA resolution images. For real-time face detection in high resolution image frames, Mustafah et al. proposed two high resolution (5 MP) smart camera-based systems by extending the works of [9-10]. Cheng et al. also proposed a face detection method in high resolution imagery by employing a grid sampling and a skin color based approach [11]. For dealing with too low-resolution faces in surveillance videos in [3] and [12], networks of two active cameras were proposed to detect faces in a wide-view camera and later extract high resolution face images in a narrow-view camera by using PTZ control.

When a face is detected in a video frame, instead of detecting face in further frames of that video clip it can be tracked. Tracking of a face in video frames has been proposed by using still camera and active PTZ camera in [13] and [14-15], respectively. In [14], an adaptive algorithm was employed along with some features extracted from face motion in order to calculate pan and tilt parameters (not including zoom parameter) of the camera to track a face. In [15], the authors however proposed a general object tracking framework by using a multi-step feature evaluation process and applied it to a face tracking application. They further extended their method to extract high resolution face sequences [7].

In addition to the above mentioned face detection and face tracking methods in video, a number of methods proposed for face quality assessment and/or face logging. Kamal et al. proposed a face quality assessment system in video sequences by using four quality metrics- resolution, brightness, sharpness, and pose [4]. This method was further extended for near infrared scenario and was included two more quality metrics- eye status (open/close) and mouth status (open/closed) [16]. In [17-18], two FQA methods have been proposed in order to improve face recognition performance. Instead of using threshold based quality metrics, [17] used a multi-layer perceptron neural network with a face recognition method and a training database. The neural network learns effective face features from the training database and checks these features from the experimental faces to detect qualified candidates for face recognition. On the other hand, Wong et al. used a multi-step procedure with some probabilistic features to detect qualified faces [18]. Kamal et al. explicitly addressed posterity facial logging problem by building sequences of increasing quality face images from a video sequence [2]. They employed a method which uses a fuzzy combination of primitive quality measures instead of a linear combination. This method was further improved in [19] by incorporating multi-target tracking capability along with a multi-pose face detection method.

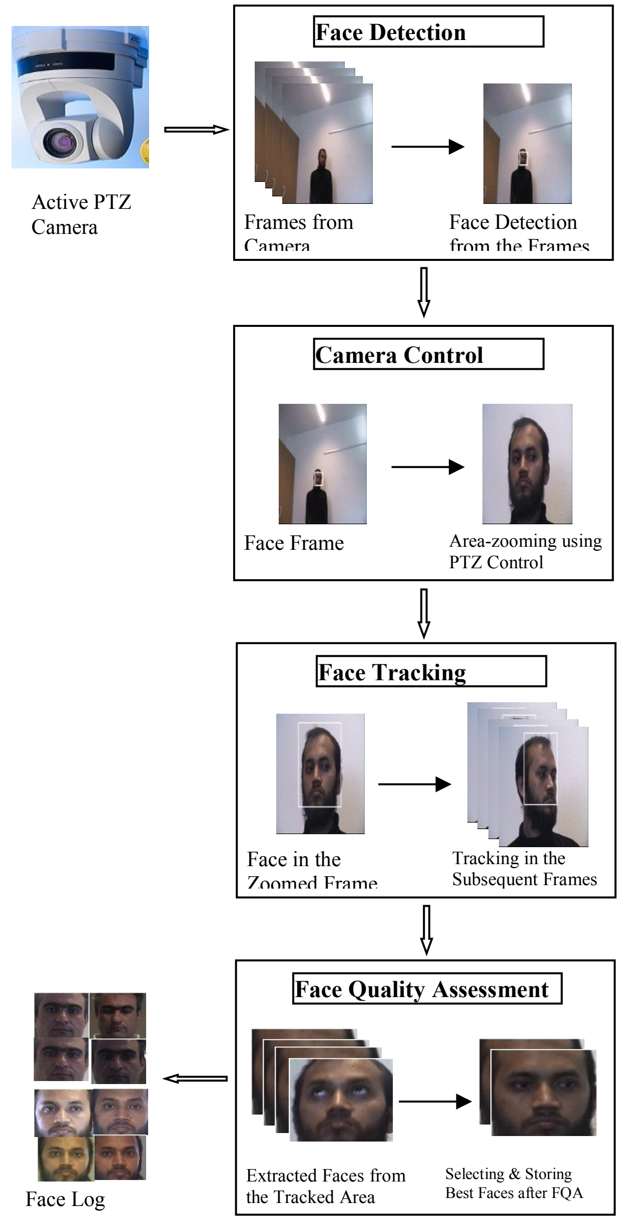

Figure 2. Block diagram of the proposed facial image acquisition system with face quality assessment

Through our intensive study so far, we observed that networks of wide-view and narrow-view PTZ camera have been employed for high resolution facial image capture [3, 7, 12]. However, the quality of images was not assessed while the PTZ features of the cameras are active. On the other hand, some FQA methods have been utilized for posterity logging of face sequences from offline video clips [2, 4, 17-18]. Thus, in this article, we propose a way to real-time capturing of high quality facial image sequences by addressing both quality and time complexity issues in an active PTZ camera capture.

The rest of the paper is organized as follows. Section II presents the proposed approach. Section III describes the experimental environment and results. Section IV concludes the paper.

2. The Proposed Approach

A typical single PTZ camera based facial image acquisition system consists of a face detection module, a camera control module, a face tracking module, and a face logging module [7]. Face detection module detects face in an image while camera sensor set in wide angle view. Camera control module utilizes the PTZ features of the camera in order to zoom in narrowly the face area of the image. Face tracking module tracks the face in subsequent frames captured by the camera. Finally, face logging module stores the faces for further processing. While logging the faces, low quality faces in terms of resolution, sharpness, brightness, and pose can be stored. Moreover, non-face images can also be logged due to false positives produced by tracking module. Thus, FQA module can be employed before face logging module. However, all these steps are computationally expensive and need to be solved real-time in order to develop an active camera based acquisition system. This paper proposes a real-time acquisition of high quality facial image from an active PTZ camera by assessing some quality metrics using a FQA method.

The architecture of the proposed system are depicted in Figure 2 and described in the following subsections.

2.1. Face Detection Module

This module is responsible to detect face in the image frames. A well-known face detection approach has been employed from [8], which is known as Viola and Jones method. This method utilizes so called Haar-like features extracted by using Haar wavelet. The Haar-like features are applied to sub-windows of the gray-scale version of an image frame with varying scale and translation. An integral image is used to reduce the redundant computation involved in computing many Haar-like features across an image. A linear combination of some weak classifiers is used to form a strong classifier in order to classify face and non-face by using an adaptive boosting method. In order to speed up the detection process an evolutionary pruning method from [20] is employed to form strong classifiers using fewer classifiers. In the implementation of this article, the face detector was empirically configured using the following constants:

- Minimum search window size:
  - 10x10 pixels in the initial camera frames
  - 70x70 pixels in the zoomed and tracked frames
- The scale change step: 10% increase per iteration
- Merging factor: 3 overlapping detections

After initializing the camera, face detection module continuously run and try to detect face. Once the face is detected, it transfers the control to the camera control module by sending face ROI of the frame.

2.2. Camera Control Module

When a face is detected in the face detection module, the size of the face is determined by the face bounding rectangle obtained from the output of the face detection module. Suppose, an input image frame is expressed by a 2-tuples *I(h, w)*, where *h* is the height of the image in pixels and *w* is the width of the image in pixels. Then, a face ROI in that image frame can be expressed by a 4-tuples *F(x, y, h, w)*, where *(x, y)* is the upper-left point of the rectangle in the image frame, and *h* and *w* are the height and width of the face, respectively. Camera control module utilizes this information from the face detection module and automatically controls the camera to zoom the face.

An efficient camera control strategy has been defined in [15] and inspired by this we utilized a concept of area zooming in order to control the camera by using PTZ parameters. Area zooming works by combining one pan-tilt command and one zoom command into one package and send it to the camera via a HTTP (Hyper-Text Transfer Protocol) command sending sub-module in the implementation. This command package is expressed by a 3-tuples *AreaZoom(ptX, ptY, Z)* with the following semantics:

AreaZoom ->
{
ptX, x-coordinate of the center of zoomed frame
ptY, y-coordinate of the center of zoomed frame
Z, scale of zooming
}

When we have a face ROI *F(x, y, h, w)* in a frame, the parameters of the AreaZoom can be calculated as follows:

$$ptX = x + \mathrm{int}(h/2) \quad (1)$$

$$ptY = y + \mathrm{int}(w/2) \quad (2)$$

$$Z = \mathrm{int}((f_{\min}/(h*w))*100) \quad (3)$$

Where, *int(..)* casts a floating point value to an integer and f_{min} is the minimum expected area size of the face after zooming. When camera receives *AreaZoom* command, it zooms in the face ROI by taking the face center point as the center point of the zoomed frame. However, in practice, zooming process takes some time to be executed. Thus, a rapidly moving face can be out of the expected zoomed frame, while camera is actually zooming in. In order to overcome this problem, we go through a face redetection phase. If there is no face, the camera is set back to the initial position and the system resumes from the initial face detection position.

2.3. Face Tracking Module

Instead of detecting face in each video frame by employing a computationally expensive face detection algorithm in full-size frames, a face tracker can be employed. When the camera control module zooms the face ROI and re-detects face in the zoomed frame, the control is transferred to the tracking module. Tracking module tracks the face and gives the most exactly matching candidate region as the face region in the subsequent zoomed frames by employing a tracking algorithm. Later, face detection algorithm is applied on the candidate region instead of full image frame to detect the face. However, the tracking algorithm should be computationally inexpensive in order to ensure real-time acquisition of face image. Among different tracking algorithms, we select a computationally inexpensive and highly competent tracking method known as *Camshift*. The performance of *Camshift* is evident from [19] and [21].

Camshift tracker takes the input image in HSV color model, finds the object center by iteratively examining a search window in the image frame by employing color histogram technique, and finally detects the window size and rotation for the object in the current frame. The readers are suggested to see [19] and [21] to get a detailed description of *Camshift*. When *Camshift* tracker selects a candidate region as the tracked face and a face detector validates it as face, the face is extracted from the image frame. Later, it is transferred to the FQA module in order to measure the quality metrics.

2.4. Face Quality Assessment Module

FQA module is responsible to assess the quality of the extracted faces from the video sequences captured by the camera. Which parameters can effectively determine face quality have been thoroughly studied in [2]. They selected some parameters among a number of parameters defined by International Civil Aviation Organization (ICAO) for identification documents. A short list of important parameters is shown in Table 1. Among these parameters four parameters were selected for real-time camera capturing and face log generation. These parameters are: out-of-plan face rotation (pose), sharpness, brightness, and resolution. In this paper, we utilize these four features in order to determine the face quality. A normalized score with the range *[0:1]* is calculated for each parameter by employing empirical thresholds for the parameters and a linear combination of scores has been utilized to generate single score for each face. Thresholds are determined by using a scaling factor with the analysis of [2]. Best faces are selected by thresholding the final single score for each face. The basic calculation process of the parameters is described in [2], we, however, include a short introduction below for readers' interest. We also include the changes necessary in the equations below in order to make it compatible with our proposed system to run in online real-time mode.

No.	*Parameter Name*	*ICAO Requirement*
1	Image Resolution	At least 420x525
2	Sharpness & Brightness	Not specified
3	Horizontal Eye Position	Center
4	Vertical Eye Position	50-70% of the image height
5	Head Width	Between 4:7 and 1:2
6	Head Height	70-80% of image height
7	Head Rotation	About 5 degrees

Table 1 Some Significant Face Quality Assessment Parameters Defined by ICAO [2].

- Pose estimation - Least out-of-plan rotated face: The face ROI is first converted into a binary image and the center of mass is calculated using:

$$x_m = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} ib(i,j)}{A}, \quad y_m = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} jb(i,j)}{A} \tag{4}$$

  Then the geometric center of face region is detected and the distance between the center of region and center of mass is calculated by:

$$Dist = \sqrt{(x_c - x_m)^2 + (y_c - y_m)^2} \tag{5}$$

  Finally the normalized score is calculated by:

$$P_{Pose} = \frac{Dist_{Th\_\max} - Dist}{Dist_{Th\_\max} - Dist_{Th\_\min}} \tag{6}$$

  Where (x_m, y_m) is the center of mass, b is the binary face image, m is the width, n is the height, A is the area of image, x_1, x_2 and y_1, y_2 are the boundary coordinate of the face, and $Dist_{Th\_max}$ and $Dist_{Th\_min}$ are the delimiters of the allowable pose values.

- Sharpness: Sharpness of a face image can be affected by motion blur or unfocused capture:

$$Sharp = abs(A(x,y) - lowA(x,y)) \tag{7}$$

  Sharpness's associated score is calculated:

$$P_{Sharp} = \frac{Sharp - Sharp_{Th\_\min}}{Sharp_{Th\_\max} - Sharp_{Th\_\min}} \tag{8}$$

  Where, $lowA(x,y)$ is the low-pass filtered counterpart of the image $A(x,y)$, and $Sharp_{Th\_max}$ and $Sharp_{Th\_min}$ are the delimiters of the allowable sharpness to be accepted.

- Brightness: This parameter measures whether a face image is too dark to use. It is calculated by the average value of the illumination component of all pixels in an image. Thus, the brightness score is calculated by (9), where $I(i,j)$ is the intensity of pixels in the face image, and $Bright_{Th\_max}$ and $Bright_{Th\_min}$ are the delimiters of the allowable brightness to be qualified.

$$P_{Bright} = \frac{\left(\sum_{i=1}^{n} \sum_{j=1}^{m} I(i,j) \Big/ (m*n) \right) - Bright_{Th\_min}}{Bright_{Th\_max} - Bright_{Th\_min}} \quad (9)$$

- Image size or resolution: Depending upon the application, face images with higher resolution generally yield better results than lower resolution faces. The score for image resolution is calculated by (10), where w is image width, h is image height, $Width_{th}$ and $Height_{th}$ are two thresholds for expected face height and width, respectively.

$$P_{Size} = \min\left\{1, \frac{w}{Width_{th}} \times \frac{h}{Height_{th}}\right\} \quad (10)$$

The final single score for each face is calculated by linearly combining the abovementioned 4 quality parameters with empirically assigned weight factor, as shown in (11):

$$Quality_{Score} = \frac{\sum_{i=1}^{4} w_i P_i}{\sum_{i=1}^{4} w_i} \quad (11)$$

Where, w_i are the weight associated with P_i, and P_i are the score values for the parameters pose, sharpness, brightness and resolution, consecutively.

Finally, the faces exceeding a quality threshold $Quality_{Th}$ are selected for logging from the captured imagery.

3. Experimental Results

3.1. Experimental Environment

In order to implement the system in an experimental environment, we used an off-the-shelf Axis PTZ 214 IP camera. The camera specification is given in Table 2. The camera was connected with the computer via an Ethernet switch and camera control was accomplished by HTTP commands. The underlying algorithms of the system was implemented in Visual C++ environment by utilizing two libraries OpenCV (computer vision) and POCO (network package).

The empirical threshold values used in the implementation are listed in Table 3. These threshold values were determined by considering the quality metrics requirements listed in Table 1 and the analysis showed in [2, 4]. We recorded several video sequences by using the proposed camera capture system, which involved both male and female objects from indoor and outdoor environment in different lighting conditions. Faces were extracted from the frames of 8 video clips (named as, *Sq1, Sq2, Sq3, Sq4, Sq5, Sq6, Sq7,* and *Sq8*) and annotated to high-quality and low-quality faces by a human observer. This annotated dataset further used to evaluate the performance.

Parameter Name	***Specification***
Angle of View, FPS	2.7^0 - 48^0 , 21-30/s
Shutter Time and Resolution	1/10000s to 1s, 0.4MP
Pan Range and Speed	$\pm$ 170^0 range, 100^0/s speed
Titl Range and Speed	-30^0 – 90^0 range, 90^0/s speed
Zoom	18x optical, 12x digital

Table 2 Camera Specification for Experimental Setup.

Parameter Name(s)	***Value(s)***
$Dist_{Th\_max}$, $Dist_{Th\_min}$	*95, 30*
$Sharp_{Th\_max}$, $Sharp_{Th\_min}$	*105, 30*
$Bright_{Th\_max}$, $Bright_{Th\_min}$	*150, 80*
$Width_{Th}$, $Height_{Th}$	*310, 275*
$Quality_{Th}$	*0.65*
Weights: w1, w2, w3, w4	*0.9, 1, 0.7, 0.5*

Table 3 Parameters Used in the Implementation.

3.2. Performance Evaluation

We employed the proposed approach to extract high quality faces from camera frames. Figure 3 depicts some example faces extracted from *Sq1, Sq2, Sq3,* and *Sq4* by the system using FQA. Figure 4 shows some example faces which were discarded due to failure to meet the quality requirements. As *Camshift* sometimes track face-colored non-face region and generate false-positive results as shown in Figure 4(a), face re-detection after *Camshift* tracking plays important role to reduce this error.

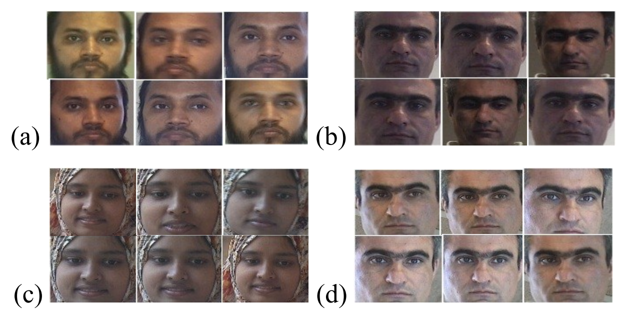

Figure 3. Some examples of extrated high-quality face from four recorded video clips: (a) *Sq1,* (b) *Sq2,* (c) *Sq3,* and (d) *Sq4*

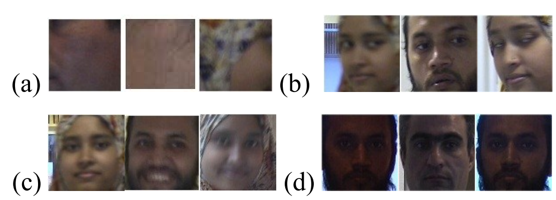

Figure 4. Some examples of discarded faces due to failure of meeting quality requirements: (a) *Not face,* (b) *Pose,* (c) *Sharpness*, and (d) *Brightness*

Table 4 summarizes the performance analysis data of the proposed approach on the aforementioned 8 annotated video clips. The clips contain 9857 image frames, including 550 faces. From the results it is observed that, in comparison with human perception, the proposed approach effectively determine the quality of the extracted faces with 85.27% accuracy. The recall rate (measure of positive cases can be caught by the system) and precision rate (measure of correct positive prediction) for positive faces are 85.30% and 98.24%, respectively. Beside these measures, the results of the proposed system for computational expenses are also impressive. The proposed system records the experimental dataset and logs good quality faces in an average of 16.20 fps (frames per second). This includes the time required for camera moving operations too. The average time required to process each frame by employing face detection, tracking and quality assessment is about 61.74 ms.

Seq.	*No. of Frames*	*No. of Tracked Frame*	*No. of Faces in Tracked Frame*	*Correct Detection TP/TN*	*False Detection FP/FN*	*Processing Time (s)*	*Accuracy & Speed*
Sq1	1479	890	120	49/49	19/5	91	Accuracy 85.27%
Sq2	1669	683	49	37/4	0/8	103	
Sq3	656	201	48	43/1	1/3	40	Recall 85.30%
Sq4	1491	813	18	13/4	1/0	94	
Sq5	2037	1103	82	51/17	4/10	124	Precision 98.24%
Sq6	447	188	71	52/7	2/10	28	
Sq7	1034	770	124	48/65	2/9	64	Speed 16.20 fps
Sq8	1044	660	38	3/26	3/6	65	

Table 4 Performance Analysis of High Quality Face Acquisition on the Experimental Dataset.

4. Conclusions

In this paper, an active camera based high-quality real-time facial image acquisition system was proposed. The quality of the faces was measured in terms of face resolution, pose, brightness, and sharpness. Only good quality faces were logged on by discarding the non-qualified imagery. Experimental result showed that the system is doable for real-time applications with an overall accuracy of 85.27% while compared with human annotated data. However, the face detection framework (Viola & Jones approach) utilized by the system is not robust to high degree of pose variation and color histogram based *Camshift* tracking often produce false positive results. Therefore, an effort is necessary to improve the system by addressing these issues. Not to mention, incorporating multi-agent tracking within real-time framework of an active camera is also an important issue to be addressed.

References

[1] S. Z. Li, and A. K. Jain, Handbook of Face Recognition, 2nd Edition, DOI 10.1007/978-0-85729-932-1, 2011.

[2] K. Nasrollahi, and T. B. Moeslund, "Complete face logs for video sequences using face quality measures," *IET Signal Porcessing*, vol. 3, no. 4. Pp. 289-300, DOI 10.1049/iet-spr.2008.0172, 2009.

[3] S. J. D. Prince, J. Elder, Y. Hou, M. Sizinstev, and E Olevskiy, "Towards Face Recognition at a Distance," *Proc. Of the IET Conf. on Crime and Security*, pp. 570-575, 2006.

[4] K. Nasrollahi, and T. B. Moeslund, "Face Quality Assessment System in Video Sequences," *1st European Workshop on Biometrics and Identity Management, Springer-LNCS,* vol. 5372, pp. 10-18, 2008.

[5] J. Tian, and K. K Ma, "A survery on super-resolution imaging," *Signal, Image and Video Processing*, vol. 5, no. 3, pp. 329-342, 2011.

[6] P. Milanfar, Super-Resolution Imaging, ISBN: 13:978-1-4398-1931-9, 2010.

[7] T. B. Dinh, N. Vo, and G. Medioni, "High Resolution Face Sequences from a PTZ Network Camera," *Proc. of the IEEE Int. Conf. on Automatic Face & Gesture Recognition and Workshops*, pp. 531-538, 2011.

[8] P. Viola, and M. Jones, "Robust real-time face detection," *Proc. Of the 8th IEEE Int. Conf. on Computer Vision*, pp. 747, 2001.

[9] Y. M. Mustafah, T. Shan, A. W. Azman, A Bigdeli, and B. C. Lovell, "Real-Time Face Detection and Tracking for High Resolution Smart Camera System," *Proc. Of the 9th Biennial Conf. of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications*, pp. 387-393, 2007.

[10] Y. M. Mustafah, A. Bigdeli, A. W. Azman, and B. C. Lovell, "Face Detection System Design for Real-Time High Resolution Smart Camera," *Proc. Of the 3rd ACM/IEEE Int. Conf. on Distributed Smart Cameras*, pp. 1-6, 2009.

[11] X. Cheng, R. Lakemond, C. Fookes, and S. Sridharan, "Efficient Real-Time Face Detection For High Resolution Surveillance Applications," *Proc. Of the 6th Int. Conf. on Signal Proce. and Comm. Systems,* pp.1-6, 2012.

[12] F. W. Wheeler, R. L. Weiss, and P. H. Tu, "Face Recognition at a Distance System for Surveillance Applications," *Proc. Of the 4th IEEE Int. Conf. on Biometrics*, pp. 1-8, 2010.

[13] F. Nicolo, S. Parupati, V. Kulathumani, and N. A. Schmid, "Near real-time face detection and recognition using a wireless camera network," *Proc. of Signal Processing, Sensor Fusion, and Target Recognition XXI*, vol. 8392, DOI 10.1117/12.920696, 2012.

[14] P. Corcoran, E. Steingerg, P. Bigioi, and A. Brimbarean, "Real-Time Face Tracking in a Digital Image Acquisition Device," *European Patent No. EP 2 052 349 B1*, pages 25, 2007.

[15] P. S. Dhillon, "Robust Real-Time Face Tracking Using an Active Camera," *Advances in Intellignet and Soft Computing: Computational Intelligences in Security for Information Systems*, vol. 63, pp. 179-186, 2009.

[16] T. Dinh, Q. Yu, and G. Medioni, "Real Time Tracking using an Active Pan-Tilt-Zoom network Camera," *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 3786-3793, 2009.

[17] J. Long, and S. Li, "Near Infrared Face Image Quality Assessment System of Video Sequences", *Proc. of the 6th Int. Conf. on Image and Graphics*, pp. 275-279, 2011.

[18] Y. Wong, S. Chen, S. Mau, C. Sanderson, and B. C. Lovell, "Patch-based Probabilistic Image Quality Assessment for Face Selection and Improved Video-based Face Recognition," *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, pp. 74-81, 2011.

[19] A. D. Bagdanov, and A. D. Bimbo, "Posterity Logging of Face Imagery for Video Surveillance," *IEEE MultiMedia*, vol. 19, no. 4, pp. 48-59, 2012.

[20] J. Jun-Su, and K. Jong-Hwan, "Fast and Robust face Detection using Evolutionary Prunning," *IEEE Trans. On Evolutionary Computation*, vol. 12, no. 5, pp. 562-571, 2008.

[21] A. Salhi, and A. Y. Jammoussi, "Object tracking system using Camshift, Meanshift and Kalman filter," *World Academy of Science, Engineering and Technology*, vol. 64, pp. 674-679, 2012.